Michigan Center for Biological Information Statewide Computational Biology Infrastructure

Athey, Brian D.*, Jagadish, H.V., Bliton, A.C., Boyd, Andrew D., Kostov, Georgi, Ogden, Jeffrey, de Wet, Jeffrey R., Chapman, Adriane, Jayapandian, Magesh, Yu, Cong, Santos, Carlos, States, David J.

University of Michigan, Ann Arbor, MI, USA

The Michigan Center for Biological Information (MCBI) is a collaborative center of Biomedical Informatics and Computational Biology for the State of Michigan. MCBI is part of Michigan's Core Technology Alliance, including University of Michigan, Wayne State University, Michigan State University, and the Van Andel Institute; and has a working group at each institution. MCBI is implementing multiple strategies for networked data integration, data sharing, and integration of high productivity genomics, proteomics, and structural biology constituents. Two approaches are being pursued towards integration of biological databases; the Michigan Molecular Interactions project (MiMI), and the Michigan Database Mirror (MIDMIR) project.

In the MiMI project, new techniques for deep data integration have been developed to incorporate multiple types of experimental data into a single flexible data model. Multiple sources of information on bio-molecular interactions have been combined into a single database. The full proteome interaction map for *Campylobacter jejuni*, currently being created by the Michigan Proteome Consortium, is being added to and correlated with the information in the MiMI database. Key features of the MiMI data model include provenance tracking and reliability metrics for every individual attribute value.

In the MIDMIR project, biological databases have been conceptually organized according to their biological relationships, and those databases most critical to our research community have been chosen to be mirrored in a database farm at the University of Michigan center. Additionally, Ensembl/Jemboss mirrors are located on-site and at MCBI nodes in Michigan, and copies of data are ported to MCBI's Biosys1 computing cluster for analysis with a set of PISE-interface bioinformatics tools, mpiBLAST, and user algorithms. The PISE-interfaces on the cluster and the Jemboss mirrors will also be used as platforms for algorithm publishing and sharing on MCBI's resources among Michigan's bioinformatics community.

A unique configuration of the compute cluster was required to allow both interactive use via the iNquiry bioinformatics portal (a commercial implementation of PISE tools by the BioTeam) and large-job scheduling via the command line, using the Sun Grid Engine scheduler and by partitioning the cluster for fair sharing between these two core user groups. MCBI has currently added a second computing cluster at our Michigan State University Node that will be used mainly for structural biology applications but also for genomics and proteomics tasks. A current activity is establishing a GRID computing environment between MCBI's clusters at UM and MSU, and other UM resources at its Center for Advanced Computing, through UM's MGRID project, with links eventually into Wayne State University's GRID activity. By application of these distributed approaches, the MCBI is making the MiMI and MIDMIR databases and associated computing capabilities available to researchers across the state of Michigan. For more information, see http://www.ctaalliance.org/MCBI/ .